

Early Adoption of the ESMF by the Land Information System

Sujay Kumar
Goddard Earth Sciences
and Technology Center,
UMBC, Baltimore, MD 21250

Paul Houser
Hydrological Sciences
Branch, NASA GSFC,
Greenbelt, MD, 20771

Christa Peters-Lidard
Hydrological Sciences
Branch, NASA GSFC,
Greenbelt, MD, 20771

Abstract—Knowledge of land surface water, energy, and carbon conditions are of critical importance due to their impact on many real world applications such as agricultural production, water resource management, and flood, weather, and climate prediction. The need to predict accurate, optimal fields of land surface states and fluxes globally has led to the development of the Land Information System (LIS), a near-real-time, high resolution system that integrates satellite and ground-based observational data along with advanced land modeling and data assimilation models. LIS employs the use of scalable, high performance computing technologies to meet the increased computational and memory requirements. LIS is adopting various Earth system modeling standards such as the Earth System Modeling Framework (ESMF) and Assistance for Land Modeling Activities (ALMA) for future coupling with other earth system models. By providing coupling capabilities through the adoption of ESMF, LIS allows the reuse of underlying high performance computing tools and land modeling infrastructure for highly scalable coupled applications.

I. INTRODUCTION

Land surface modeling seeks to predict the terrestrial water, energy, and biogeochemical processes by solving the governing equations of soil-vegetation-snowpack medium. The land surface and atmosphere are coupled to each other over a variety of time scales through the exchanges of water, energy, and carbon. An accurate representation of land surface processes is critical for improving models of the boundary layer and land-atmosphere coupling at all spatial and temporal scales and over heterogeneous domains. Long term descriptions of land use and fluxes also enable in the accurate assessments of climate characteristics. In addition to the impact on the atmosphere, predicting land surface processes is also critical for ecosystem modeling and water resources prediction and management.

The need for accurate predictions of land surface states has led to the development of a Land Data Assimilation System (LDAS) [4] that consists of a number of uncoupled land models using remotely-sensed and in-situ observations within a land data assimilation framework. LDAS has been successfully used for simulations up to 1/8 degree resolution in both real-time and long-term (50 years) retrospective simulations. However, to improve the understanding of land surface processes and their interaction with atmospheric processes, such a system needs to be implemented globally at a high resolution such as 1km. The motivation behind LIS is to extend the capabilities of LDAS and to provide a computing environment for global, high resolution land surface modeling.

Many existing Earth science applications, though highly scalable and computationally capable, lack the ability to interoperate with other Earth system applications. The cost to adapt such applications to function with other Earth system applications may be prohibitively high. LIS attempts to achieve code interoperability by applying advanced software engineering concepts in its design. The system is designed in a modular fashion with different modules capturing the behavior of various system components. The use of object oriented principles helps in designing LIS to be flexible and extensible, enabling rapid prototyping of new applications into LIS.

In addition to providing an infrastructure to support land surface research and applications activities, LIS is adopting other Earth system modeling standards and conventions, such as ESMF [3] and ALMA [1]. ESMF is a system that provides a flexible software infrastructure to foster interoperability, portability and code reuse in climate, numerical weather prediction, data assimilation and other Earth science applications. ALMA is a land-atmosphere coupling standard that is being developed by the broad land-atmosphere research community. By conforming to the standards laid out by ESMF and ALMA, LIS can provide capabilities to interact with other Earth system models.

The following sections describe the land modeling and computing infrastructure, the interoperable features, and the adoption of ESMF and ALMA standards, tools and interfaces in LIS.

II. LAND SURFACE MODELING IN LIS

LIS includes a model control and input/output system that drives multiple offline one-dimensional land surface models (LSMs) to facilitate global modeling. The one-dimensional LSMs apply the governing equations of the physical processes of the soil-vegetation-snowpack medium to characterize the transfer of mass, energy, and momentum between a vegetated surface and the atmosphere.

The global land surface is modeled by dividing it into two-dimensional regions or gridcells (for example cells of size $1\text{km} \times 1\text{km}$ would lead to approximately 5×10^8 gridcells). Assuming approximately 15ms for each day of land surface model execution on a particular gridcell, it can be estimated that to conduct a day's simulation at 1km using 15 minute timesteps would require approximately 3 months. It is clear

that global, high resolution land surface modeling presents a significant computational challenge.

In addition to the computational challenges, high resolution land surface modeling also presents significant I/O bottlenecks. The definition of data structures associated with each gridcell and the variables associated with defining sub-grid variability at the global resolution causes the memory requirements to scale significantly with the domain resolution. From the experiments conducted at low resolutions using LDAS, it can be estimated that memory in the order of terabytes would be required for global land surface modeling at 1km.

Due to the computational challenges described above, the use of scalable computing technologies is critically important and relevant for LIS. Land surface processes have rather weak horizontal coupling on short time and large space scales. LIS is taking advantage of this inherent parallelism that enables highly efficient scaling across massively parallel computational resources. A unique custom-built 200 node Linux cluster consisting of 192 computing nodes and 8 I/O nodes is employed by LIS to handle the huge data management requirements of high resolution land surface modeling. LIS currently includes high performance capabilities to perform global land surface modeling up to 5km global resolution.

III. INTEROPERABILITY IN LIS

LIS system is designed using object oriented design principles, providing a number of well-defined interfaces or “hook points” for enabling development of new features and applications into LIS. As described earlier, LIS incorporates a number of one- dimensional LSMs. These LSMs typically require three types of inputs: 1) initial conditions, which describe the initial state of land surface; 2) boundary conditions, which describe both the atmospheric states also known as forcings and the soil states; and 3) parameters, which are functions of soil, vegetation, topography, etc. and are used to solve the governing equations. LIS makes use of data from various satellite and ground-based observations as well as output from numerical prediction models to force the LSMs in LIS. A common interface to different LSMs enable the reuse of this broad set of data and other tools. Further, forcing different LSMs with the same input allows researchers to perform intercomparisons of output from different LSMs. The need for such a flexible interface is addressed in LIS by a component-based design that provides a number of well defined interfaces for the incorporation of new land surface schemes. The implementation of these interfaces allows the user to employ underlying computing and land modeling infrastructure when a new LSM is incorporated.

LIS will comply with ALMA standards to interoperate with other land modeling systems. ALMA is a flexible data exchange convention to facilitate the exchange of forcing data for LSMs and the results produced by them. The output data variables and formats, and the variables passed between LIS and the land models follow the ALMA specification. By implementing the ALMA convention, LIS can exchange data with other land modeling systems that are also ALMA

compliant. Further, ALMA compliance enables LIS to be used for intercomparisons of land surface models.

IV. LIS AND ESMF

The purpose of ESMF is to develop a framework that provides a structured collection of building blocks that can be customized to develop model components. ESMF can be broadly viewed as consisting of an infrastructure of utilities and data structures for building model components and a superstructure for coupling and running them. The use of ESMF interfaces and utilities in LIS allows for future coupling with earth system models such as atmospheric models.

A. ESMF Infrastructure Adoption

The ESMF infrastructure layer contains both higher level data handling objects and lower level utility routines. The infrastructure layer provides abstractions for fields and group of fields discretized on grids in classes such as *Field*, *Grid*, *Bundle* etc. LIS plans to adopt these representations necessary for implementing coupled applications and to make use of some of the infrastructure utilities such as *Regrid*.

The utility layer presents a uniform interface for common system functions such as time manager, basic communications, error handler, diagnostics, etc. LIS currently uses the ESMF time manager and the logging and error diagnostics tools. The time management utility provides useful functions for time and data calculations and higher level functions that control model time stepping and alarms. The log utility organizes diagnostic output and allows for searches and filters to be constructed. The error handler provides both uniform handling of errors and a way for users to select how the errors will be handled. Figure 1 shows a schematic view of the interaction between LIS and ESMF infrastructure. The solid lines represent the utilities currently used by LIS and the dotted lines represent the some of the tools that will be implemented in future.

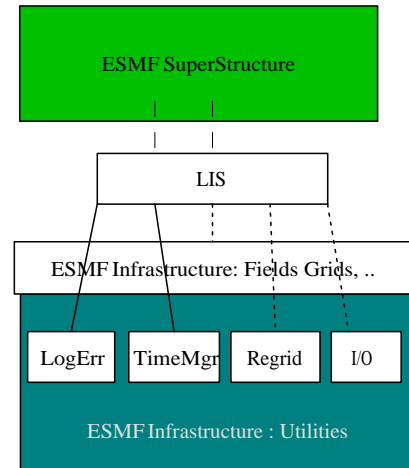


Fig. 1. LIS and ESMF

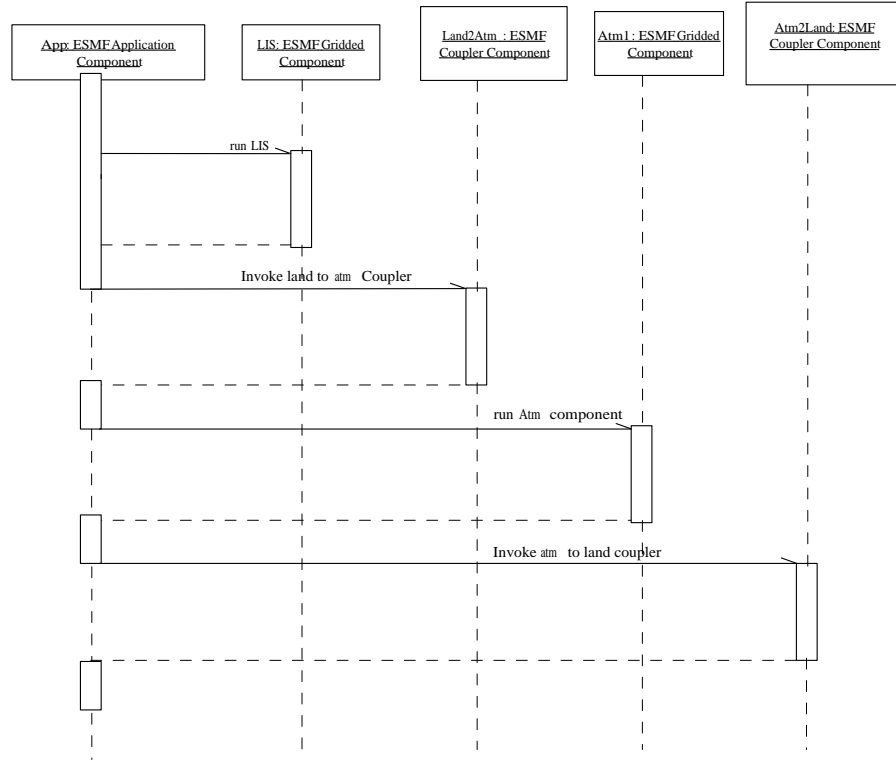


Fig. 2. Sequence diagram for running a simple coupled application using LIS as an ESMF Gridded Component

B. ESMF Superstructure Adoption

ESMF also defines a number of guidelines for applications that are intended to be coupled with other Earth system models. ESMF provides definitions of a Gridded Component class for user-supplied components discretized on grids and a Coupler Component class for the software that is used to couple them together. LIS will implement the interfaces required to be a Gridded Component and will use ESMF_State class to exchange information with other models and systems. A land surface model can be coupled with other earth system models by implementing it as a Gridded Component. However, since LIS provides the infrastructure to drive different offline land surface models, implementing LIS itself as a Gridded Component will allow any LSM in LIS to be used for coupling with other earth system models. LIS could serve as the land modeling component in the coupled system, providing the best possible surface fluxes to the atmospheric modeling components. Figure 2 shows a simple sequence diagram for running an application with LIS being coupled to an atmospheric model, exchanging data through custom-defined couplers.

V. SUMMARY AND FUTURE DIRECTIONS

LIS is an evolving system for global high resolution land surface modeling. Some of the computational challenges faced by LIS are addressed by the use of scalable parallel computing architectures as described in the results above. However, the I/O requirements still pose a significant challenge and will

be the focus of the development efforts in the future. LIS is also actively participating in the development of Earth system modeling standards by complying with the ESMF guidelines and standards. ESMF is a project in development and LIS will adopt the utilities and interfaces as they become available. In addition to adopting the ESMF for future coupling with other earth system models, LIS itself provides interoperable tools for the land modeling community by adopting land surface modeling standards such as ALMA and by defining extensible interfaces in its design.

REFERENCES

- [1] ALMA. <http://www.lmd.jussieu.fr/alma/>.
- [2] CLM. <http://www.cgd.ucar.edu/tss/clm/>.
- [3] ESMF. <http://www.esmf.ucar.edu>.
- [4] LDAS. <http://ldas.gsfc.nasa.gov>.
- [5] NOAH. <ftp://ftp.ncep.noaa.gov/pub/gcp/ldas/noahslm/>.